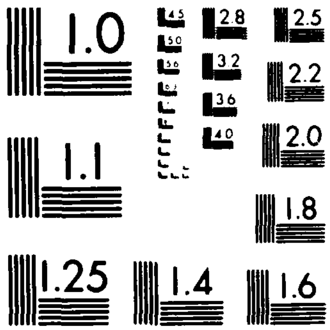
AD-A186 043 PREDICTING MAGAZINE AUDIENCES WITH A LOGLINEAR MODEL 1/1
(U) FLORIDA STATE UNIV TALLAHASSEE DEPT OF STATISTICS
P J DANAHER JUL 87 FSU-TR-M758 AFOSR-TR-87-1082
UNCLASSIFIED F49620-85-C-0007 F/G 12/3 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

AD-A186 043

REPORT DOCUMENTATION PAGE

DTIC FILE COPY

1b. RESTRICTIVE MARKINGS	
2a. SECURITY CLASSIFICATION AUTHORITY	N/A
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	N/A
3. DISTRIBUTION/AVAILABILITY OF REPORT	Approved for Public Release; Distribution Unlimited
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	FSU Technical Report No. M -758
5. MONITORING ORGANIZATION REPORT NUMBER(S)	AFOSR-TR- 87-1082
6a. NAME OF PERFORMING ORGANIZATION	Florida State Universeity
6b. OFFICE SYMBOL (If applicable)	
7a. NAME OF MONITORING ORGANIZATION	AFOSR/NM
6c. ADDRESS (City, State and ZIP Code)	Department of Statistics Tallahassee, FL 32306-3033
7b. ADDRESS (City, State and ZIP Code)	Building 410 Bolling AFB, DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION	AFOSR
8b. OFFICE SYMBOL (If applicable)	NM
9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER	F49620-85-C-0007 AFOSR ~~85-C-0007~~
8c. ADDRESS (City, State and ZIP Code)	Building 410 Bolling AFB, DC 20332-6448

10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
6.1102F	2304	A5	

11. TITLE (Include Security Classification)	Predicting Magazine Audiences with a Loglinear Model
12. PERSONAL AUTHOR(S)	Peter J. Danaher
13a. TYPE OF REPORT	~~Technical~~ Journal
13b. TIME COVERED	FROM ____ TO ____
14. DATE OF REPORT (Yr., Mo., Day)	July 1987
15. PAGE COUNT	27
16. SUPPLEMENTARY NOTATION	

DTIC ELECTE SEP 2 9 1987 S D E

17. COSATI CODES		
FIELD	GROUP	SUB. GR.

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

A loglinear model for predicting magazine exposure distributions is developed and it's parameters are estimated by using the maximum likelihood technique. The accuracy of the loglinear and a Dirichlet-multinomial model are compared using 1985 AGB: McNair data. The results show that the loglinear model has significantly smaller prediction errors than the Dirichlet-multinomial model. A simple algorithm for optimal media scheduling is given.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐
21. ABSTRACT SECURITY CLASSIFICATION	UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL	Maj Crowley
22b. TELEPHONE NUMBER (Include Area Code)	202-767-5025
22c. OFFICE SYMBOL	AFOSR/NM

DD FORM 1473, 83 APR

EDITION OF 1 JAN 73 IS OBSOLETE.

SECURITY CLASSIFICATION OF THIS PAGE

AFOSR-TR- 87-1082

Predicting Magazine Audiences with a Loglinear Model

by

Peter J. Danaher

Department of Statistics
Florida State University
Tallahassee, Florida 32306

Accession For	
NTIS GRA&I	X
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

QUALITY INSPECTED 2

FSU Technical Report M758
AFOSR Technical Report 87-207

July 1987

Research supported by the Air Force Office of Scientific Research Grant AFOSR 85-C-0007.

87 8 25 009

PREDICTING MAGAZINE AUDIENCES WITH A LOGLINEAR MODEL

by

Peter J. Danaher*

Department of Statistics
Florida State University
Tallahassee, FL 32306
ph. (904) 644-3578

*This paper is taken from the author's 1987 Florida State University Ph.D. dissertation for which the author gratefully acknowledges the assistance of Professor Duane A. Meeter. The author also wishes to thank AGB: McNair Surveys New Zealand Ltd. for providing the data used in this study.

Currently the author is a graduate student in the Statistics Department at Florida State University.

PREDICTING MAGAZINE AUDIENCES WITH A LOGLINEAR MODEL

Abstract

A loglinear model for predicting magazine exposure distributions is developed and it's parameters are estimated by using the maximum likelihood technique. The accuracy of the loglinear and a Dirichlet-multinomial model are compared using 1985 AGB: McNair data. The results show that the loglinear model has significantly smaller prediction errors than the Dirichlet-multinomial model. A simple algorithm for optimal media scheduling is given.

Introduction

The annual advertising budget for Anheuser-Busch Incorporated in 1983 was $290,616,400 (*Advertising Age* 1984). Having spent such an enormous amount of money on advertising, the company wants to be certain their money is being spent efficiently. Their advertising money is not spent efficiently if they have poor estimates of the potential audience for their advertising campaigns. An underestimate of the audience does not reflect the actual efficiency of the advertising campaign; instead, a reduced expenditure may still achieve the advertiser's desired audience. An improvement in the audience estimates for Anheuser-Busch, reducing the annual advertising budget by only 1%, would amount to a saving of $2,906,164.

Most ad agencies and advertisers use the *exposure distribution* (e.d.) to assess campaign efficiency. The e.d. is the proportion of the target population which sees none, one, two, or up to all of the ads in the campaign. In turn, the e.d. is used to estimate *reach*, the proportion of the population which is exposed to at least one insertion, *frequency*, the proportion of the population which is exposed to at least one, at least two, or up to all of the ads, and *effective reach*, the mean of the e.d. Effective reach is also known as "Gross Rating Points" (Naples 1979).

Many methods have been employed to estimate the e.d. However, the majority are *ad hoc* and lack accuracy over a range of advertising campaigns (Chandon 1976). A survey of media directors from 288 advertising agencies in the U.S. showed that media directors place "more sophisticated reach analyses" high on their list of important future media research services (Russell and Martin 1980). This finding was corroborated by a more recent survey (Leckenby and Kishi 1982b) which showed that 90% of media directors from the 100 largest advertising agencies believe that between "some" and a "great deal" of improvement is needed in e.d. models. The same survey indicated that only about one-third of the directors perceive that the models they currently use estimate the observed e.d. to within five percent.

Another important use of e.d. estimates is in media selection (Aaker 1975; Lee 1962, 1963; Little and Lodish 1969). All advertising campaigns have a budget. It usually costs more to advertise in magazines with high readership. An advertiser can achieve a high reach by placing many ads in a magazine with a high readership although the campaign cost will be high also. Alternatively it is usually possible to achieve high reach by placing ads in a range of magazines, with more ads in lower readership magazines, thereby reducing the overall campaign cost. Accurate reach estimates are needed to juggle insertion placement in a variety of magazines so as to maximize the reach or effective reach, whilst keeping within the budget.

The first known attempt to estimate reach for a media schedule was by Agostini (1961). However, Agostini's method was not successful when applied to British magazines (Metheringham 1964). Metheringham's (1964) own method used the beta-binomial distribution (BBD) (Skellam 1948; Ishii and Hayakawa 1960). The BBD has also been used as an e.d. model by Greene and Stock (1967), Liebman and Lee (1974) and Chandon (1976). In addition the BBD has been applied to TV schedules (Headen, Klompmaker and Teel 1977; Rust and Klompmaker 1981), consumer purchasing behavior (Chatfield and Good-

hardt 1970; Morrison 1979), household disease distributions (Griffiths 1973), proportions with extraneous variance (Kleiman 1973) and as an indicator of TV loyalty (Sabavala and Morrison 1977). However, Metheringham's method was shown to overestimate reach and odd-numbered exposures (Schreiber 1969; Chandon 1976). His method also fails to estimate the characteristic shape of the e.d. (Liebman and Lee 1974; Leckenby and Kishi 1982a).

Chandon (1976) made an empirical comparison of six *ad hoc*, ten stochastic and three simulation methods of estimating the e.d. His study showed that two of the better methods used Waring's formula (cf. Feller 1969) to estimate terms in the formula where triplicate and higher order probabilities were needed. One method was introduced by Kwerel (1964) and the other by Hofmans (1966). An empirical study by Leckenby and Kishi (1982a) showed that Hofmans' method is superior to Kwerel's.

A model currently popular for estimating the e.d. is based on the Dirichlet- multinomial distribution (DMD) (Chandon 1976; Leckenby 1981; Leckenby and Kishi 1982a, 1984). The DMD has also been applied to brand choice models (Goodhardt, Ehrenberg and Chatfield 1984), mixed media models (Rust and Leone 1982, 1984) and to pollen counts (Mosimann 1962). The best known method to date for estimating the e.d. is a DMD model of Leckenby and Kishi's (1984) which uses Hofman's geometric distribution to estimate the between-vehicle duplication. However, the major deficiency with Leckenby and Kishi's model is that it is designed primarily for schedules with equal insertions in all the magazines, but equal insertion placement infrequently occurs in practice. Leckenby and Kishi state that their model can be used for unequal insertions by "treating multiple insertions in a vehicle as single insertions in multiple identical vehicles." The danger with this procedure is that it ignores the within-vehicle duplication, a factor known to be important due to reading loyalty.

We propose to develop three models to allow for e.d.s with one, two, and three or more

magazines. The models build on each other in that the model for one magazine is used to improve the fit of the model for two magazines and the model for two magazines is used to estimate the parameters of the model for three or more magazines. In contrast to some past e.d. models we will use rigorous statistical methodology in developing the models for magazine e.d.s in addition to employing statistically-efficient parameter estimation.

The model for one magazine introduces an additional parameter to the BBD to account for reading loyalty. This simple generalization of the BBD was first suggested by Chandon (1976) although his method of parameter estimation was naive, sometimes leading to inconsistencies in his estimates. We will improve upon his estimation method.

The model for two magazines uses the DMD to model a bivariate magazine e.d. The model is capable of handling unequal insertions and, again, we will use efficient parameter estimation.

The model for three or more magazines uses a loglinear model to estimate the cells of a multidimensional contingency table. Bishop, Fienberg and Holland (1975) give a full account of the derivation and application potential for loglinear models. We make a reasonable simplifying assumption for the loglinear model from which it follows that only bivariate data are needed to fit the model. Thus our model, for it's improved accuracy, uses only as much data as existing models.

Our three models are empirically tested against the best currently-known models using sample exposure data from the AGB:McNair "National Media Survey" conducted in 1985. Finally we give a simple algorithm for which our e.d. models can be used to maximize either reach or effective reach while keeping within a campaign budget.

The Models

One-Magazine Model

Many people subscribe to one or more magazines. Among them is a proportion which always reads a particular magazine.

Chandon (1976) suggested a modification of the BBD which he called the "two segment beta-binomial model". One segment is definite readers and the other is probable readers. They represent proportions ω and 1-ω respectively. The parameter ω may be viewed as a loyalty factor. A high value of ω indicates an appreciable reading loyalty whilst a low value indicates little or no loyalty to a particular magazine.

By mixing this loyalty proportion with the mass function of the BBD we obtain the modified BBD (MBBD). Let X be the number of exposures a person has to k insertions in a single magazine. The mass function of the MBBD is

$$(1) \qquad f^{MB}(X=x) = (1-\omega)\binom{k}{x}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta+k)}\frac{\Gamma(k-x+\beta)}{\Gamma(\beta)}\frac{\Gamma(x+\alpha)}{\Gamma(\alpha)} + \omega \mathrm{I}_{\{X=k\}}\ , \qquad x = 0, \ldots, k,$$

where $\Gamma(l+1) = l\Gamma(l)$, the usual gamma fuction, and $\mathrm{I}_{\{X=k\}}$ is 1 when $X = k$ and 0 otherwise. Note that when ω=0 the MBBD reduces to the BBD.

Chandon (1976) estimated α, β and ω by equating the sample proportion of non-readers to the proportion of nonreaders given by the MBBD model for $k = 1, 2, 3$. These data cannot be obtained from our sample data (cf. Q1 and Q2 below). Chandon's system of three equations sometimes results in an inconsistent solution which forces him to set $\omega = 0$, thereby losing any advantage of using the MBBD.

We use maximum likelihood estimation to estimate α, β and ω. The likelihood equations are given in the technical appendix. Danaher (1987) proved that the maximum likelihood estimates (MLEs) of α, β and ω are best asymptotically normal (BAN), which means that as the sample size (n) tends to infinity these estimates tend to their true values and these estimates have the smallest variance among all estimates. Serfling (1980, p. 142) gives a formal definition of this concept.

Two-Magazine Model

Let Y_1 = the number of exposures *exclusive* to magazine 1, Y_2 = the number of exposures *exclusive* to magazine 2, and Y_3 = the number of exposures to *both* magazines 1

and 2, with $0 \leq Y_i \leq k$, $i = 1,2,3$, i.e., initially there are k insertions in both magazines.

Let $\mathbf{Y}=(Y_1, Y_2, Y_3)$; then the DMD has mass function (Mosimann 1962)

$$f^{DM}(\mathbf{Y}=\mathbf{y}) = \frac{k!}{(k-\sum_{i=1}^{3} y_i)!}\frac{\Gamma(\tau)}{\Gamma(\tau+k)}\frac{\Gamma(k-\sum_{i=1}^{3} y_i+\gamma_0)}{\Gamma(\gamma_0)}\prod_{i=1}^{3}\frac{\Gamma(y_i+\gamma_i)}{\Gamma(\gamma_i)y_i!},$$

$$\gamma_i > 0,\ \tau = \sum_{i=0}^{3}\gamma_i,\ 0 \leq y_i \leq k,\ i=1,2,3,\ \sum_{i=1}^{3} y_i \leq k.$$

Now let $X_1 = Y_1 + Y_3$ and $X_2 = Y_2 + Y_3$; then (X_1, X_2) is the bivariate e.d. of the numbers of exposures to magazines 1 and 2 with mass function,

$$\begin{aligned} g_{X_1,X_2}(X_1=x_1, X_2=x_2) = {} & \frac{k!\Gamma(\tau)}{\Gamma(\tau+k)}\sum_{x_3=max\{0,x_1+x_2-k\}}^{min\{x_1,x_2\}} \\ & \frac{\Gamma(x_1-x_3+\gamma_1)\Gamma(x_2-x_3+\gamma_2)\Gamma(x_3+\gamma_3)\Gamma(k+x_3-x_1-x_2+\gamma_0)}{(x_1-x_3)!(x_2-x_3)!x_3!(k+x_3-x_1-x_2)!\prod_{i=0}^{3}\Gamma(\gamma_i)}, \\ & 0 \leq x_i \leq k,\quad i=1,2. \end{aligned} \tag{2}$$

At this stage we have a model for a bivariate e.d. when there are equal insertions in both magazines. If the media schedule requires unequal insertions then add dummy insertions to the magazine with the lesser insertions to get insertion equality. This method of adding dummy insertions was used by Rust and Leone (1984) for two different media, viz., TV and magazines. Following the construction of the bivariate e.d. using (2) a hypergeometric adjustment is made (Chandon 1976; Rust and Leone 1984) as follows. Suppose k_i insertions are placed in magazine i, $i = 1,2$. When $k_1 > k_2$,

$$h_{X_1^*,X_2^*}(x_1^*, x_2^*) = \sum_{z=x_2^*}^{k_1-k_2+x_2^*}\frac{\binom{k_2}{x_2^*}\binom{k_1-k_2}{z-x_2^*}}{\binom{k_1}{z}}g_{X_1,X_2}(x_1^*, z)\ ,\quad 0 \leq x_i^* \leq k_i\ ,\ i=1,2. \tag{3}$$

Now $h_{X_1^*,X_2^*}$ is the bivariate e.d. model when the insertions are unequal and g_{X_1,X_2} comes from (2).

The final step in constructing the bivariate e.d model is to adjust (3) so that it has MBBD marginals from (1) by using iterated proportions (Deming and Stephan 1940).

To estimate the parameters of (3) we will use the data from Q1 below where $k = 1$ and denote it $\{n_{x_1 x_2}\}$, where $n_{x_1 x_2}$ is the number of people in the sample who have x_1 exposures to magazine i, $i = 1,2$ $(0 \leq x_i \leq 1)$, and n is the sample size. The problem now is that we have only three linearly independent data (since $n = \sum_{x_1,x_2} n_{x_1 x_2}$) and four parameters to estimate.

The MLEs of the γ_i are

$$\text{(4)} \qquad \hat{\gamma}_0 = \frac{n_{00}}{n}\hat{\tau}, \; \hat{\gamma}_1 = \frac{n_{10}}{n}\hat{\tau}, \; \hat{\gamma}_2 = \frac{n_{01}}{n}\hat{\tau}, \; \hat{\gamma}_3 = \frac{n_{11}}{n}\hat{\tau} \, .$$

From (4) it can be seen that in order to estimate the γ_i, $i = 0,1,2,3$, it is necessary to estimate τ also. Basu and de B. Pereira (1982) prove that X_1 and X_2 have beta-binomial distributions with parameters α_i and β_i defined as follows

$$\text{(5)} \qquad \begin{aligned} \gamma_1 + \gamma_3 = \alpha_1 \, , \quad & \gamma_0 + \gamma_2 = \beta_1 , \\ \gamma_2 + \gamma_3 = \alpha_2 \, , \quad & \gamma_0 + \gamma_1 = \beta_2 \, . \end{aligned}$$

Let $\hat{\alpha}_i$ and $\hat{\beta}_i$ be MLEs or method of moments estimates of α_i and β_i, $i = 1,2$. From (5), $\alpha_i + \beta_i = \gamma_0 + \gamma_1 + \gamma_2 + \gamma_3 = \tau$, $i = 1,2$. However, in general $\hat{\alpha}_1 + \hat{\beta}_1 \neq \hat{\alpha}_2 + \hat{\beta}_2$ so that some authors take a weighted average of $\hat{\alpha}_1 + \hat{\beta}_1$ and $\hat{\alpha}_2 + \hat{\beta}_2$ to estimate τ with $\frac{\sum_{i=1}^{2} w_i(\hat{\alpha}_i + \hat{\beta}_i)}{w_1 + w_2}$, where $w_i = \frac{\hat{\alpha}_i}{\hat{\alpha}_i + \hat{\beta}_i}$, $i = 1,2$ (Chandon 1976, Leckenby and Kishi 1984, Rust and Leone 1984). We found this unappealing since this estimator of τ is rather *ad hoc*. The authors could equally well have chosen the arithmetic, geometric, or harmonic mean of $(\hat{\alpha}_i + \hat{\beta}_i)$, $i = 1,2$. In addition this estimate is usually inconsistent, i.e., as the sample size tends to infinity, the weighted average estimate of τ above does not tend to τ (Danaher 1987).

Our estimate of τ is

$$\text{(6)} \qquad \hat{\tau} = n \left(\frac{\hat{\alpha}_1 \hat{\beta}_1 \hat{\alpha}_2 \hat{\beta}_2}{(n_{10} + n_{11})(n_{00} + n_{01})(n_{01} + n_{11})(n_{00} + n_{10})} \right)^{\frac{1}{4}} \, .$$

This estimate is numerically close to the geometric mean of $(\hat{\alpha}_i + \hat{\beta}_i)$, $i = 1,2$. This estimate is consistent and after (3) is adjusted to conform to MBBD marginals, by iterated

proportional fitting, the estimate of the bivariate e.d. obtained by substituting (4) and (6) into (3) is BAN (see Danaher 1987).

Model for Three or More Magazines

Here we construct a loglinear model to estimate the m-dimensional probabilities for an m-variate e.d. To fix ideas we will examine the case $m = 3$ and estimate the 3-dimensional e.d. of (X_1, X_2, X_3), where X_j is the number of exposures a person has to magazine j, $j = 1, 2, 3$.

Using the notation of Fienberg (1977) the general loglinear model for three dimensions is

$$\log\, m_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{12(ij)} + u_{13(ik)} + u_{23(jk)} + u_{123(ijk)}\ , \qquad i = 0, 1 \ldots k_1,\ j = 0, 1, \ldots k_2,\ k = 0, 1, \ldots k_3\ , \tag{7}$$

where m_{ijk} = the expected number of people in the sample exposed to i out of k_1 insertions in magazine 1, j out of k_2 insertions in magazine 2 and k out of k_3 insertions in magazine 3. Constraints on the parameters in (7) are given in Fienberg (1977).

Let x_{ijk} = the number of people in the sample exposed to i, j and k insertions in magazines 1, 2 and 3 respectively.

The loglinear model with no second-order interaction has $u_{123(ijk)} = 0$ for all i, j, k in (7). We tested the validity of the no second-order interaction assumption on 220 combinations of three magazines with one insertion in each magazine. Using an asymptotic chi-squared test given by Feinberg (1977 33) we found only 15 of the 220 combinations had significant second-order interaction at the 5 % level of significance. Since we have a large sample size of 5201 the asymptotic chi-squared test is powerful which means we can be confident that the assumption of no second-order interaction is reasonable for most of our data.

There is no closed-form solution for the MLEs of $\{m_{ijk}\}$ but Fienberg (1977), provides an iterative solution, performing the following steps to obtain the MLEs of $\{m_{ijk}\}$.

Step 1: set $\hat{m}_{ijk}^{(0)} = 1$ for all i, j, k, then for $\nu = 0$,

Step 2: $\hat{m}_{ijk}^{(3\nu+1)} = \frac{x_{ij+}}{\hat{m}_{ij+}^{(3\nu)}} \hat{m}_{ijk}^{(3\nu)}$,

Step 2: $\hat{m}_{ijk}^{(3\nu+2)} = \frac{x_{i+k}}{\hat{m}_{i+k}^{(3\nu+1)}} \hat{m}_{ijk}^{(3\nu+1)}$,

Step 4: $\hat{m}_{ijk}^{(3\nu+3)} = \frac{x_{+jk}}{\hat{m}_{+jk}^{(3\nu+2)}} \hat{m}_{ijk}^{(3\nu+2)}$,

where $x_{ij+} = \sum_k x_{ijk}$ etc.

Now repeat steps 2 to 4 for $\nu = 1, 2, \ldots$ until the change in $\hat{m}_{ijk}$ *'s* from one cycle to the next is small. This technique is called iterative proportional fitting and was introduced by Deming and Stephan (1940). We have already used this technique for the two-magazine model.

It can be seen from steps 2 to 4 above that the only data required to fit the loglinear model with no second-order interaction is $\{x_{ij+}\}$, $\{x_{i+k}\}$ and $\{x_{+jk}\}$. In a typical situation for a loglinear fit the $\{x_{ijk}\}$ come from the sample (Fienberg 1977). It is possible to get $\{x_{ijk}\}$ from the sample data for some values of $i, j,$ and k (c.f. Q1 and Q2 below). However, if we knew $\{x_{ijk}\}$ for all i, j and k there would be no need to *estimate* the e.d.

Since we only need to know $\{x_{ij+}\}$, $\{x_{i+k}\}$ and $\{x_{+jk}\}$ the sensible thing to do is obtain accurate estimates of these three sets of bivariate statistical-frequencies. Hence, our method is to pretend that our two-magazine bivariate e.d. estimates obtained from (3) (and marginal adjustment) are the observed values in steps 2 to 4 above. Note that we must multiply the probability estimates by n to get the "observed values".

Once we have $\{\hat{m}_{ijk}\}$ we can estimate the mass function of the three dimensional e.d. (with no second-order interaction). Now let $X = \sum_{j=1}^{3} X_j$ be the total number of exposures an individual has to a media schedule. Then the mass function of X (denoted $f(X)$) is estimated by

$$\text{(9)} \qquad \hat{f}(X = x) = \frac{1}{n} \sum_{\{(i,j,k):i+j+k=x\}} \hat{m}_{ijk} \,, \qquad x = 0, 1, \ldots, \sum_{j=1}^{3} k_j.$$

An efficient algorithm was written to compute the $\{\hat{m}_{ijk}\}$, then construct (9). It is

clear how this procedure, with the assumption of no second-order or higher interaction generalises to m magazines ($m \geq 3$).

Danaher (1987) proved that the estimates of the e.d. probabilities arising from this loglinear model are BAN.

We can think of the two-magazine DMD model above as being a "saturated" loglinear model (Fienberg 1977). Hence, from now on we will consider that we have two models, the MBBD for one magazine and a loglinear model for two or more magazines.

Empirical Tests

The Available Data

To fit and test the accuracy of our e.d. model we used the AGB:McNair Surveys New Zealand Ltd. "National Media Survey" data from a sample of $n = 5201$ residents of New Zealand, 10 years or older. This multistage cluster sample covered the period July through December 1985, with about 200 people interviewed each week. It is assumed that reading habits are stationary over time so that we may think of all 5201 people as a sample from a single population. Many questions were asked of the respondents although the only two question formats relevant to us were (for weekly magazines),

Q1) "Have you personally read or looked into any issue of ...(magazine name) in the last seven days - it doesn't matter where?" (Has a Y/N answer)

Q2) "How many different issues of ...(magazine name) do you personally read or look into in an average month - it doesn't matter where?" (Has answer 0,1,2,3,4 issues)

Table 1 shows how Q1 and Q2 are modified for monthly, two-weekly, and two-monthly magazines. Q1 and Q2 were asked for 40 different magazines. There were ten general appeal magazines, eight business magazines, six were women's magazines, six were news magazines, five were hobbies magazines and five were sports oriented. The magazines ranged in single issue reach from 0.7% to 39% of New Zealand's population.

We extract univariate exposure data from Q2 and bivariate exposure data from Q1,

Table 1: Question Inserts for Different Magazine Periods.

Magazine	Q1 period	Q2 period
Weekly	last seven days	average month
Two-weekly	last two weeks	last three months
Monthly	last month	last six months
Two-monthly	last two months	last twelve months

to fit the models above. The Simmon's data used by Leckenby and Kishi (1984) allowed them to test their model only for schedules with up to two insertions in each magazine. In contrast, we can test our models at insertion levels ranging from 1 (using Q1), 4 (for weekly magazines, using Q2) and 6 (for nonweeklies, using Q2). We can also test our models for schedules with unequal insertions by using a mixture of Q1 and Q2 to get the observed e.d. This should permit a thorough test of our models.

Test Design

Since our one-magazine model, the MBBD, is different from the loglinear model for two or more magazines we will test this model separately. The test is exhaustive in that the MBBD model is fitted then used to predict observed sample e.d.s for all 40 magazines at two insertion levels, $k = 1$ and $k = 4$ (for weeklies) or $k = 6$ (for nonweeklies), making 80 total schedules. We chose the BBD as the best currently available model to compare with MBBD.

For schedules with two or more magazines we constructed a completely randomized design, selecting 340 schedules ranging in size from 2 to 6 vehicles. The schedules were chosen so that half of them had equal insertions. Of the schedules with equal insertions half had one insertion in each magazine while the other half had 4 or 6 insertions in each magazine. For the schedules with equal insertions we also fitted Leckenby and Kishi's (1984) "DMD1" model (renamed DMDLK) as this model performed the best out of the eight models they tested.

In the case of the two-magazine schedules, with one insertion in each magazine, both the loglinear and DMDLK models reproduce the sample data exactly, thereby giving no

prediction error. Owing to this, our equal-insertion schedules for two magazines have only insertion levels 4 or 6.

Definition of errors

The most common technique for assessing model accuracy is to use the model to estimate the e.d. for a schedule whose observed sample e.d. is already known (Liebman and Lee 1974; Chandon 1976; Leckenby and Kishi 1982a, 1984; Rust and Leone 1984).

Denote $f_i = f(X = i)$ and $\hat{f}_i$, $i = 0, 1, \ldots, k$, respectively, as the observed and estimated probabilities of the e.d. Reach and effective reach are denoted ρ and ρ_e. Four measurements of error will be used. They are,

i) Mean squared error (MSE) where

$$MSE = \frac{1}{k+1} \sum_{i=0}^{k} (f_i - \hat{f}_i)^2 \; ;$$

ii) Relative error in reach (RER) where

$$RER = \frac{|\hat{f}_0 - f_0|}{1 - f_0} = \frac{|\hat{\rho} - \rho|}{\rho};$$

iii) Error in the exposure probabilities over schedule reach ($EPOR$) where

$$EPOR = \frac{\sum_{i=1}^{k} |f_i - \hat{f}_i|}{1 - f_0} \; ;$$

iv) Absolute error in effective reach ($AEER$) where

$$AEER = |\rho_e - \hat{\rho}_e| \; .$$

The MSE is a popular measure of error in statistics whilst the other three measures of error are designed more specifically for magazine exposure models. RER and $EPOR$ have been used by Liebman and Lee (1974) and Leckenby and Kishi (1982a, 1984). Notice that $(RER + EPOR)(1 - f_0) = \sum_{i=0}^{k} |f_i - \hat{f}_i|$, which is the sum of the absolute errors.

Thus $EPOR$ may be thought of as the contribution to the sum of absolute errors which is attributable to nonzero exposure levels.

Results and Discussion

The average errors for the four error measures are given in Table 2. The MBBD has smaller average error than the BBD for all but the RER. A one-way analysis of variance showed that the difference in the average errors between the two models is significant (p-value$<$ 0.025) for the $EPOR$ and $AEER$. That is, the MBBD gives significantly smaller errors than the BBD for half the error measures.

Insert Table 2 About Here

Table 2 also shows the variation of the average errors for the loglinear and DMDLK models across schedule sizes ranging from 2 to 6. The loglinear model gives smaller average errors than the DMDLK for all the error measures and schedule sizes. In addition, the loglinear model's errors do not increase in magnitude as the schedule size increases from 3 to 6, as occurs for the DMDLK. The overall average for the entire 340 schedules is given in the last row of Table 2 and in Table 3 these averages are compared by using a one-way analysis of variance. The result of the analysis of variance F-tests is that the loglinear model gives significantly smaller (p-value$<$ 0.001) errors than the DMDLK for all four error measures.

Insert Table 3 About Here

We were able to test the loglinear model for schedules with unequal insertions. The magnitude of the errors for these schedules did not differ significantly from the errors for schedules with unequal insertions.

Generally, the errors for schedules having one insertion in each magazine were smaller, by a factor of at least 10, than for schedules with unequal or high, equal, insertions. For this reason the average errors for the two-magazine schedules are unexpectedly high (being larger than for three-magazine schedules, for instance). This is because all the

two-magazine schedules had four or more insertions in each magazine, as explained above.

One drawback of the loglinear model is that it's computation time increases as schedule size and insertion levels increase. For example, schedules comprised of three and six magazines, with each magazine having four insertions, took 0.1 and 201 CPU seconds, respectively, on a CYBER 760 mainframe. However, the increased computation time is not felt to be a major shortcoming since, in practice, few schedules exceed six magazines (Stroeven, Managing Director of AGB:McNair Suveys N.Z. Ltd., personal communication). Danaher (1987) discusses ways in which the computation time for very large schedules can be significantly reduced, at the expense of accuracy.

Optimal Media Scheduling

Now that we have excellent models for estimating the e.d. we can use these models to answer the following question, "How many insertions should be placed in each magazine to maximize reach or effective reach whilst keeping within a predetermined budget?"

Suppose we have m magazines with k_j insertions in each magazine, at a cost of c_j for each insertion, $j = 1, 2, \ldots, m$. Let $f(X = x)$ denote the e.d. mass function as before and let the total allowable campaign cost be C. A formal statement of the integer programming optimization problems is to vary $(k_1, k_2, \ldots, k_m)$ so as to

$$\begin{aligned} \text{Maximize} \quad & \rho = \ 1 - f(X = 0) \quad \text{(reach)}, \\ \text{or} \quad & \rho_e = \sum_{j=1}^{m} k_j \frac{(\alpha_j + \omega_j \beta_j)}{\alpha_j + \beta_j} \quad \text{(effective reach under the MBBD model)}, \\ \text{subject to} \quad & \sum_{j=1}^{m} k_j c_j \leq C, \qquad k_j = 0, 1 \ldots < \infty, \ j = 1, \ldots, m. \end{aligned}$$

Most integer programming techniques are suited only to linear objective functions (Garfinkel and Nemhauser 1972). We can see that ρ_e is linear in k_j, but ρ is not, so we need to use some nonstandard techniques to maximize ρ subject to the constraint. Solution algorithms for this optimization problem by the branch and bound (Garfinkel

and Nemhauser 1972) and dynamic programming (Denardo 1982) methods are available. A feasible solution is an m-tuple $(k_1, \ldots, k_m)$ which satisfies the above constaint. We devised an *ad hoc* solution method which considers the entire solution space, defined by the constraint $\sum_{j=1}^{m} k_j c_j \leq C$, searching for the feasible solution with the highest reach or effective reach. However, we can eliminate many of the feasible solutions by exploiting the following monotonicity property of reach.

Denote by $\rho(k_1, \ldots, k_m)$ the reach achieved by placing k_j insertions in the j^{th} magazine. If we add one insertion to the l^{th} magazine then

$$\rho(k_1 \ldots, k_l + 1, \ldots, k_m) > \rho(k_1 \ldots, k_l, \ldots, k_m)\ , \quad l = 1, \ldots, m,\ k_j \geq 0,\ j = 1, \ldots, m.$$

In words, this inequality says that adding an insertion to any magazine is guaranteed to increase the reach.

This componentwise monotonicity property of reach is very important since it greatly reduces the computation needed to solve the reach optimization problem. For instance, we know $\rho(2,1,1) > \rho(1,1,1)$ so if (2,1,1) is a feasible solution there is no need to find $\rho(1,1,1)$ ((1,1,1) is clearly also a feasible solution) since we know it to be less than $\rho(2,1,1)$.

A formal statement of the algorithm is:

$k_j^{(0)} = \lfloor C/c_j \rfloor$, $j = 1, \ldots, m$, $\rho_{max} = 0$,
Order the magazines so that $c_1 \leq c_2 \leq \cdots \leq c_m$ then,
for j=1 to m
for $i = k_j$ to 0 step -1
if $\sum_{j=1}^{m} k_j c_j \leq C$ then
    $\rho_{max} \leftarrow max\{\rho_{max},\ \rho(k_1 \ldots, k_m)\}$
    update optimal solution if ρ_{max} changes
    $k_1 \leftarrow k_1^{(0)}$
    $k_{j+1} \leftarrow k_{j+1} - 1$
    begin j loop over again
    end if
end i loop
end j loop.

Effective reach also has the componentwise monotonicity property so the same algorithm above can be used to find the optimal schedule when maximizing effective reach.

Table 4 gives the cost and readership information for a group of three magazines as well as the optimal solution for reach and effective reach with a campaign cost, C, of \$30,000. The effective reach is maximized by putting all of the insertions in the magazine with the highest ratio of single issue reach to cost per insertion, whilst the reach is maximized by placing as many insertions as possible in the most popular magazine. There were 102 feasible solutions to this particular optimization problem of which 82 were eliminated by the the monotonicity property of reach and effective reach, an 80% reduction in the number of schedules to be tested. Without using the monotonicity property, the execution times for maximizing ρ and ρ_e are 24.4 and 0.072 CPU seconds, respectively, on a CYBER 760. If the monotonicity property is used the corresponding execution times are 4.8 and 0.061 CPU seconds. These represent, respectively, an 80% and 15% reduction in computer time. Clearly, we have made a major saving in computation time when maximizing ρ and a smaller, though significant, saving when maximizing ρ_e.

Insert Table 4 About Here

Conclusion

We used a simple generalization of the BBD suggested by Chandon (1976), viz. the MBBD, with our efficient parameter estimation and empirically showed that it performs better that the BBD.

The major contribution of this study is the application of a loglinear model to estimating multidimensional e.d.s. Leckenby and Kishi (1984) showed that exposure models based on the DMD are accurate in estimating magazine e.d.s. However, a loglinear model, with no second-order or higher interaction, is shown to be statistically significantly better than Leckenby and Kishi's most accurate DMD model. We also found that their model loses accuracy as the schedule size increases whereas the loglinear model sustains it's accuracy

over a range of schedule si... In addition, the DMDLK is not designed for schedules with unequal insertions. In contrast, the loglinear model handles unequal insertions, giving the same accuracy as for schedules with equal insertions.

The trade-off for the increased accuracy of the loglinear model is an increase in computation time. This is felt to be a tolerable shortcoming considering the significant improvement the loglinear model makes over the best currently known model.

<u>Technical Appendix</u>

Let n_i be the number of people in the sample exposed to i out of k insertions in a magazine. Denote the likelihood function for the MBBD as L and define

$$c = c(\alpha,\beta) = \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta+k)}$$

and $\Delta_1(\gamma,l) = \sum_{j=0}^{l-1} 1/(\gamma+j)$. Then the likelihood equations are

$$\begin{aligned}
\frac{\partial \log L}{\partial\alpha} &= \sum_{i=1}^{k-1}\Delta_1(\alpha,i)n_i - (n-n_k)\Delta_1(\alpha+\beta,k) \\
&\quad + \frac{(1-\omega)c[\Delta_1(\alpha,k)-\Delta_1(\alpha+\beta,k)]n_k}{\omega+(1-\omega)c}, \\
\frac{\partial \log L}{\partial\beta} &= \sum_{i=0}^{k-1}\Delta_1(\beta,k-i)n_i - (n-n_k)\Delta_1(\alpha+\beta,k) \\
&\quad - \frac{(1-\omega)c\Delta_1(\alpha+\beta,k)n_k}{\omega+(1-\omega)c}, \\
\frac{\partial \log L}{\partial\omega} &= \frac{-(n-n_k)}{1-\omega} + \frac{(1-c)n_k}{\omega+(1-\omega)c} \quad .
\end{aligned}$$

We calculate $\hat{\omega}$ by equating $\frac{\partial \log L}{\partial\omega}$ to zero and get $\hat{\omega} = \frac{n_k/n - c(\hat{\alpha},\hat{\beta})}{1-c(\hat{\alpha},\hat{\beta})}$.
When $\hat{\omega}$ is substituted into the first order partial derivatives of α and β the following second order partial derivatives result,

$$\begin{aligned}
\frac{\partial^2 \log L}{\partial\alpha^2} &= \frac{(n-n_k)c}{1-c}[(\Delta_1(\alpha,k)-(\Delta_1(\alpha+\beta,k))^2/(1-c)+\Delta_2(\alpha+\beta,k)-\Delta_2(\alpha,k)] \\
&\quad - \sum_{i=1}^{k-1}\Delta_2(\alpha,i)n_i + (n-n_k)\Delta_2(\alpha+\beta,k), \\
\frac{\partial^2 \log L}{\partial\beta\partial\alpha} &= \frac{(n-n_k)c}{1-c}[(\Delta_1(\alpha+\beta,k)-\Delta_1(\alpha,k))\Delta_1(\alpha+\beta,k)/(1-c)+\Delta_2(\alpha+\beta,k)] \\
&\quad + (n-n_k)\Delta_2(\alpha+\beta,k), \\
\frac{\partial^2 \log L}{\partial\beta^2} &= \frac{(n-n_k)c}{1-c}[(\Delta_1(\alpha+\beta,k))^2/(1-c)+\Delta_2(\alpha+\beta,k)] \\
&\quad - \sum_{i=0}^{k-1}\Delta_2(\beta,n-i)n_i + (n-n_k)\Delta_2(\alpha+\beta,k),
\end{aligned}$$

where $\Delta_2(\gamma, l) = -\frac{\partial}{\partial\gamma}\Delta_1(\gamma, l)$. The MBBD likelihood equations have no closed form solution but may be solved by the Newton-Raphson method where the above second order partial derivatives are utilized. Since $\hat{\omega}$ is an explicit function of $\hat{\alpha}$ and $\hat{\beta}$ the numerical work is considerably reduced.

Table 2: Average Errors for Schedules of size 1 to 6 vehicles: Comparison Between MBBD and BBD Models and Between Loglinear and DMDLK Models

Schedule Size	Model	Error Type $MSE \times 10^5$	$EPOR\%$	$RER\%$	$AEER\%$
1	MBBD	3.13	11.70	3.63	0.33
	BBD	5.20	15.24	3.20	0.90
2	Loglinear	4.74	14.71	1.98	1.04
	DMDLK	9.23	24.22	5.23	3.75
3	Loglinear	2.90	11.34	1.32	0.79
	DMDLK	4.90	13.48	2.24	2.68
4	Loglinear	6.98	13.89	1.74	1.91
	DMDLK	8.89	15.51	2.14	3.28
5	Loglinear	5.54	11.01	1.01	0.93
	DMDLK	18.69	18.12	2.30	4.08
6	Loglinear	4.98	11.17	1.85	1.06
	DMDLK	21.97	19.35	2.12	6.47
Overall for sizes 2-6	Loglinear	4.55	12.95	1.57	1.10
	DMDLK	8.94	18.03	3.20	3.46

Table 3: Analysis of Variance Comparison of DMDLK and Loglinear Models

Error Type	Source	d.f.	Mean Square	F
MSE	Between models	1	2.13×10^{-7}	16.2[a]
	Error	502	1.31×10^{-8}	
$EPOR$	Between models	1	0.286	27.2[a]
	Error	502	0.0105	
RER	Between models	1	0.0294	38.1[a]
	Error	502	0.00077	
$AEER$	Between models	1	0.0616	50.7[a]
	Error	502	0.00121	

[a] p-value < 0.001.

Table 4: Optimal Solution for a Three Magazine Schedule.

Magazine	Cost per Insertion $	Single Issue Reach %	Optimal Soln for ρ	Optimal Soln for ρ_e
Readers Digest	2400	25.8	0	12
Womans Weekly	5602	37.9	1	0
N.Z. Listener	6061	39.0	4	0

References

Aaker, D.A. (1975), "ADMOD:An Advertising Decision Model," *Journal of Marketing Research*, February, 37-45.

Advertising Age (1984), September 14, 55, 60.

Agostini, J.M. (1961), "How to Estimate Unduplicated Audiences," *Journal of Advertising Research*, 1, 11-14.

Basu, D. and de B. Pereira (1982), "On the Bayesian Analysis of Categorical Data: The Problem of Nonresponse," *Journal of Statistical Planning and Inference*, 6, 345-362.

Bishop, Y., Fienberg, S. and Holland, P. (1975), *Discrete Multivariate Analysis: Theory and Practice*. M.I.T. Press, Cambridge, Massachusetts.

Chandon, J-L. J. (1976), "A Comparative Study of Media Exposure Models," unpublished Ph.D. dissertation, Northwestern University.

Chatfield, C. and Goodhardt, G. (1970), "The Beta-Binomial Model for Consumer Purchasing Behavior," *Applied Statistics*, 19, 240-250.

Danaher, P.J. (1987). "Estimating Multidimensional Tables from Survey Data: Predicting Magazine Audiences," unpublished Ph.D. dissertation, Florida State University.

Deming, W.E. and Stephan, F.F. (1940), "On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals Are Known," *Annals of Mathematical Statistics*, 11, 427-444.

Denardo, E.V. (1982), *Dynamic Programming: Models and Applications*. Prentice-Hall Inc. New York.

Fienberg, S.E. (1977), *The Analysis of Cross-Classified Categorical Data*. M.I.T. Press, Cambridge, Massachusetts.

Feller, W. (1968), *An Introduction to Probability Theory and its Applications*. Vol. 1, third edition. John Wiley and Sons, New York.

Garfinkel, R.S. and Nemhauser, G.L. (1972), *Integer Programming*. John Wiley and Sons, New York.

Goodhardt, G.J., Ehrenberg, A.S.C. and Chatfield, C. (1984), "The Dirichlet as a Comprehensive Model of Buying Behavior," (with discussion). *Journal of the Royal Statistical*

Society A, 147, 5, 621-655.

Greene, J.D. and Stock, S.J. (1967), *Advertising Reach and Frequency in Magazines*. Marketmath Inc. and Readers' Digest Assoc., New York.

Griffiths, D.A. (1975), "Likelihood Estimation for the Beta-Binomial Distribution and an Application to the Household Distribution of the Total Number of Cases of a Disease," *Biometrics*, 29, 637-648.

Headen, R.S., Klompmaker, J.E. and Teel, J.E. (1977), "Predicting Audience Exposure to Spot TV Advertising Schedules," *Journal of Marketing Research*, 16, 1-9.

Hofmans, P. (1966), "Measuring the Net Cumulative Coverage of Any Combination of Media," *Journal of Marketing Research*, 3, 269-278.

Ishii, G. and Hayakawa, R. (1960), "On the Compound Binomial Distribution," *Annals of the Institute of Statistical Mathematics*, 12, 69-80.

Jeuland, A.P., Bass, F.M. and Wright, G.P. (1980), "A Multibrand Stochastic Model Compounding Heterogeneous Erlag Timing and Multinomial Choice Processes," *Operations Research*, 28, 255-277.

Kleinman, J.C. (1973), "Proportions With Extraneous Variance: Single and Independent Samples," *Journal of the American Statisttical Association*, 68, 46-54.

Kwerel, S.M. (1964), "Estimating the Unduplicated Audience of a Combination of Media Vehicles: Integrated Theory and Estimation Methods," unpublished Ph.D. dissertation, Columbia University.

Leckenby, J. (1981), "Exposure Distribution Models: Some Estimation Methods in Application," paper presented at the Annual Convention of the American Academy of Advertising, Gainsville, Florida.

__________ and Kishi, S. (1982a), "Performance of Four Exposure Distribution Models," *Journal of Advertising Research*, 22, 35-44.

__________, __________ (1982b), "How Media Directors View Reach/Frequency Estimation," *Journal of Advertising Research*, 22, 64-69.

__________, __________ (1984), "The Dirichlet-Multinomial Distribution as a Magazine Exposure Model," *Journal of Marketing Research*, 21, 100-106.

Lee, A.M. (1962), "Decision Rules for Media Scheduling: Static Campaigns," *Operational Research Quarterly*, 13, 3, 224-236.

__________ (1963), "Decision Rules for Media Scheduling: Dynamic Campaigns," *Operational Research Quarterly*, 14, 4, 365-376.

Liebman, L. and Lee, E. (1974), "Reach and Frequency Estimation Services," *Journal of Advertising Research*, 14, 23-25.

Little, J.D.C and Lodish, L.M. (1969), "A Media Planning Calculus," *Operations Research*, 17, 1, 1-35.

Metheringham, R.A. (1964), "Measuring the Net Cumulative Coverage of a Print Campaign," *Journal of Advertising Research*, 4, 23-28.

Morrison, D.G. (1979), "Purchase Intentions and Purchasing Behavior," *Journal of Marketing*, 43, 65-74.

Mosimann, J.E. (1962), "On the Multinomial Distribution, the Multivariate β- Distribution and Correlations Among Proportions," *Biometrika*, 49, 65-82.

Naples, M.J. (1979), *Effective Frequency: The Relationship Between Frequency and Advertising Awareness*. Association of National Advertisers, New York.

Russell, J.T. and Martin, C.H. (1980). "How Ad Agencies View Research," *Journal of Advertising Research*, 20, 2, 27-31.

Rust, R.T. and Klompmaker, J.E. (1981), "Improving the Estimation Procedure for the Beta-Binomial TV Exposure Model," *Journal of Marketing Research*, 18, 442-448.

__________ and Leone, R.P. (1984), "The Mixed Media Dirichlet-Multinomial Distribution: A Model for Evaluating Television-Magazine Advertising Schedules," *Journal of Marketing Research*, 21, 89-99.

Sabavala, D.J. and Morrison, D.G. (1977), "Television Show Loyalty: A Beta-Binomial Model Using Recall Data," *Journal of Advertising Research*, 17, 35-43.

Schreiber, R.J. (1969), "The Metheringham Method for Media Mix: An Evaluation," *Journal of Advertising Research*, 9, 2, 54-56.

Serfling, R.J. (1980), *Approximation Theorems of Mathematical Statistics*. John Wiley and Sons, New York.

Skellam, J.G. (1948), "A Probability Distribution Derived From the Binomial Distribution by Regarding the Probability of Success as a Variable Between Sets of Trials," *Journal of the Royal Statistical Society* B, 10, 257-265.

END

12-87

DTIC